

University of Dundee

Molecular basis for recognition of the Group A Carbohydrate backbone by the PlyC streptococcal bacteriophage endolysin

King, Harley; Ajay Castro, Sowmya; Pohane, Amol Arunrao; Scholte, Cynthia M.; Fischetti, Vincent A.; Korotkova, Natalia

Published in:
Biochemical Journal

DOI:
[10.1042/BCJ20210158](https://doi.org/10.1042/BCJ20210158)

Publication date:
2021

Licence:
CC BY

Document Version
Peer reviewed version

[Link to publication in Discovery Research Portal](#)

Citation for published version (APA):

King, H., Ajay Castro, S., Pohane, A. A., Scholte, C. M., Fischetti, V. A., Korotkova, N., Nelson, D. C., & Dorfmueller, H. C. (2021). Molecular basis for recognition of the Group A Carbohydrate backbone by the PlyC streptococcal bacteriophage endolysin. *Biochemical Journal*, 478(12), 2385-2397.
<https://doi.org/10.1042/BCJ20210158>

General rights

Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from Discovery Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

**Molecular basis for recognition of the Group A Carbohydrate backbone by the PlyC
streptococcal bacteriophage endolysin**

Harley King^{1,^}, Sowmya Ajay Castro^{2,^}, Amol Arunrao Pohane³, Cynthia M. Scholte¹,
Vincent A. Fischetti⁴, Natalia Korotkova^{3,5}, Daniel C. Nelson¹ and Helge C. Dorfmueller^{2,*}

¹Institute for Bioscience and Biotechnology Research, University of Maryland, Rockville, MD, USA

²Division of Molecular Microbiology, School of Life Sciences, University of Dundee, Dundee, UK

³Department of Microbiology, Immunology and Molecular Genetics, University of Kentucky,
Lexington, Kentucky, USA

⁴Laboratory of Bacterial Pathogenesis and Immunology, The Rockefeller University, New York,
NY, USA

⁵Department of Molecular and Cellular Biochemistry, University of Kentucky, Lexington, KY, USA

[^]equal contribution

*Correspondence: Helge C. Dorfmueller, h.c.z.dorfmueller@dundee.ac.uk

Keywords: Streptococcus; rhamnose, polysaccharide, cell wall, bacteriophage, endolysin,
pyogenes

27 **Abbreviations (alphabetical order)**

28 Cell wall-binding domain (CBD), enzymatically active domain (EAD), Glucose (Glc), glycerol
29 phosphate (GroP), Glycosyl Hydrolase (GlyH), Group A Carbohydrate (GAC), Group A
30 Streptococcus (GAS), Group A-variant Streptococcus (GAVS), Group C Carbohydrate (GCC),
31 Group C Streptococcus (GCS), Group G Carbohydrate (GGC), *N*-acetyl-glucosamine (GlcNAc),
32 *N*-acetyl-galactosamine (GalNAc), *N*-acetyl-muramic acid (MurNAc), Peptidoglycan (PG),
33 polyrhamnose (pRha), Rhamnose (Rha), *S. dysgalactiae subsp. equisimilis* (SDSE), *S. mutans*
34 serotype c carbohydrate (SCC)

35

36

Abstract

Endolysins are peptidoglycan (PG) hydrolases that function as part of the bacteriophage (phage) lytic system to release progeny phage at the end of a replication cycle. Notably, endolysins alone can produce lysis without phage infection, which offers an attractive alternative to traditional antibiotics. Endolysins from phage that infect Gram-positive bacterial hosts contain at least one enzymatically active domain (EAD) responsible for hydrolysis of PG bonds and a cell wall binding domain (CBD) that binds a cell wall epitope, such as a surface carbohydrate, providing some degree of specificity for the endolysin. Whilst the EADs typically cluster into conserved mechanistic classes with well-defined active sites, relatively little is known about the nature of the CBDs and only a few binding epitopes for CBDs have been elucidated. The major cell wall components of many streptococci are the polysaccharides that contain the polyrhamnose (pRha) backbone modified with species-specific and serotype-specific glycosyl side chains. In this report, using molecular genetics, microscopy, flow cytometry and lytic activity assays, we demonstrate the interaction of PlyCB, the CBD subunit of the streptococcal PlyC endolysin, with the pRha backbone of the cell wall polysaccharides, Group A Carbohydrate (GAC) and serotype *c*-specific carbohydrate (SCC) expressed by the Group A Streptococcus and *Streptococcus mutans*, respectively.

Introduction

Endolysins are bacteriophage-encoded PG hydrolases that normally function from within the cell to lyse the bacterial host, releasing progeny phage and completing the phage lifecycle [1]. However, the lytic activity of endolysins can be harnessed for antimicrobial use due to their ability to equally lyse bacteria when applied exogenously, without infection by a parental phage. Due to their direct lytic action on target PG, endolysins are not affected by efflux pumps, alterations in metabolism, or other mechanisms of antibiotic resistance, making them ideal candidates for development against multi-drug resistant organisms [2-4]. Notably, at least three endolysins, some of which are active against methicillin-resistant *Staphylococcus aureus*, are currently being evaluated in human clinical trials for their antimicrobial activity (reviewed in [5]).

Most endolysins, and in particular those from phage that infect Gram-positive bacterial hosts, are comprised of modular domains. An enzymatically active domain (EAD) is generally found in the N-terminal region, while a cell wall-binding domain (CBD) is located in the C-terminal region [6]. As the name implies, the EAD is a catalytic domain that is responsible for cleaving specific bonds in the PG, the nature of which is dependent on the mechanistic class of the EAD. Occasionally, endolysins contain two EADs, although both are not necessarily active. The CBD binds at high affinity [7] to a cell wall-specific epitope and was suggested to dictate genus, species and serovar-specificity of the endolysin. The CBD targets may be surface carbohydrates, wall teichoic acids linked to the Gram-positive bacterial cell wall, or the PG itself [8].

The endolysin now known as PlyC is one of the first described endolysins and remains one of the most studied. In 1934, Alice Evans noted a “nascent lysis” activity derived from streptococcal phage lysates on streptococcal strains that were not sensitive to the phage itself [9]. By 1957, Krause had determined that the phage used by Evans was specific for Group C Streptococci (GCS), but an “enzyme” produced by the phage could lyse Groups A, A-variant, and C

Streptococci (GAS, GAVS, and GCS, respectively) [10]. These findings were confirmed by Maxted later the same year and extended to include Group E Streptococci (GES) [11]. In 2001, PlyC, then referred to as the streptococcal C₁ lysin, became the first endolysin to be used therapeutically, protecting mice from GAS challenge in a nasopharyngeal model [2]. Subsequent studies revealed that PlyC is a structurally unconventional endolysin, which is not encoded by a single gene as found for all other endolysins described to date. Rather, PlyC is a nine-subunit holoenzyme encoded by two distinct genes, *plyCB* and *plyCA*, within a polycistronic operon [12]. Eight PlyCB subunits self-assemble into a ring structure and form the basis of the CBD that binds the streptococcal surface. A single PlyCA subunit contains two distinct EADs separated by an extended α -helical linker region, which interfaces with the N-terminal residues of the PlyCB octamer [13]. The PlyCA EADs consist of a glycosyl hydrolase (GH) domain and a cysteine, histidine-dependent aminohydrolase/peptidase (CHAP) domain. The very high lytic activity of PlyC relative to other endolysins is attributed to synergy/cooperativity between the two EADs.

Although the EADs of PlyCA have been extensively characterized with respect to specificity of PG bonds they cleave, active-site residues, and their synergistic activity, it is unclear how PlyCB recognizes PG of GAS, GCS and GES. Similar to other Gram-positive bacteria, the plasma membrane of streptococci is surrounded by thick cell wall that consists of a complex network of PG with covalently attached polysaccharides. Rebecca Lancefield utilized the unique immunogenicity of the surface carbohydrates in β -hemolytic streptococci to subsequently separate them into serogroups [14]. *S. pyogenes* is categorized as Group A Streptococcus, whilst *S. dysgalactiae subsp equisimilis* (SDSE) produce at least two types of carbohydrates and are annotated as Group C and G Streptococci, respectively [15]. Pioneering work by many researchers have revealed that the cell wall polysaccharides of GAS, GCS and *S. mutans*, consist of a pRha backbone modified with species-specific and serotype-specific glycosyl side chains (Fig. 1A) [16, 17]. In GAS, GCS and *S. mutans* serotype c, the polysaccharides termed the

Lancefield group A carbohydrate (GAC), group C carbohydrate (GCC), and SCC, respectively, have a conserved repeating $\rightarrow 3)\alpha\text{-Rha}(1\rightarrow 2)\alpha\text{-Rha}(1\rightarrow$ di-saccharide backbone [16, 17]. The β -*N*-acetylglucosamine (GlcNAc) side chains are attached to the 3-position of the α -1,2 linked rhamnose (Rha) in GAC [16, 18, 19]. The GCC side chains have two *N*-acetyl-galactosamine (GalNAc) residues attached to the same hydroxyl of Rha [16, 18, 20]. SCC carries the α -glucose (Glc) side chains attached to the 2-position of the α -1,3 linked Rha in SCC [17]. Additionally, the side chains of GAC and SCC are decorated in parts with the glycerol phosphate (GroP) moiety [21]. The GAC and SCC biosynthetic pathways are encoded by the 12-gene loci *gacABCDEFGHIJKL* and *sccABCDEFGHLMNPQ*, respectively. The molecular mechanisms of GAC and SCC biosynthesis have been the focus in a number of recent studies [21-25].

Remarkably, extracted GAC is known to partially inhibit the lytic actions of PlyC [26]. Furthermore, a GAS mutant carrying the unmodified GAC (lacking the GlcNAc side chains) displayed enhanced susceptibility to PlyC [25], implicating PlyCB in recognition of the pRha backbone. Here, we identify the pRha backbone of GAC as the definite minimal binding epitope for PlyCB.

Materials and Methods

Bacterial strains, growth conditions and media

Streptococcus pyogenes strain D471 was propagated on solid media in plates containing Todd-Hewitt broth supplemented with yeast extract (0.2%) (THY) and agar (1.4%) or in liquid THY as described by Gera et al [27]. *S. mutans* wild type (Xc) and mutants were grown in Todd-Hewitt broth with 1% yeast extract. All cultures were grown without antibiotics and without aeration at 37°C. *E. coli* genotypes DH5 α (NEB, cat. No. C2988J), DH10 α (Thermo Fisher, cat No. 18297010), BL21 (NEB, cat No. C2530H) and Origami 2 (Millipore Sigma, cat. No. 71346) were used for routine plasmid propagation or protein expression and grown in Lysogeny Broth (LB)

medium supplemented with either 50 µg/ml kanamycin, 100 µg/ml ampicillin or 35 µg/ml chloramphenicol as needed.

Bacterial strains *E. coli* CS2775 were transformed with pRGP1 plasmid [28], *gacABCDEFG* [23] to produce pRha or empty plasmid control (pHD0131). The bacterial cells were grown overnight in LB containing erythromycin (150 µg/ml) at 37°C and used next day for whole cell Western blots and FACS and microscopy analysis.

Recombinant expression and purification of PlyCB_{WT} and PlyCB_{R66E}

PlyCB_{WT} (GenBank ID: NC_004814.1:7517-7735) was expressed from a pBAD24 vector and purified from BL21 cells as previously described [12]. In brief, the culture was grown in LB and induced with 0.25% L-arabinose at OD₆₀₀ ~1.2-1.4 (Alfa Aesar, cat no. A11921). Cultures were grown at 37°C with shaking at 180 RPM for 3-4 hours, centrifuged at 4500 x g and resuspended in phosphate-buffered saline (PBS). Lysis was performed by a French press (1800 psi). Benzonase (Millipore Sigma, cat. No. 70746-3) was added and the lysate incubated at room temperature with rotation for 20-30 min. The lysate was centrifuged at 20,000 x g for 20 min and the cleared lysate was passed through a 0.45-µm filter and loaded onto a XK-26/20 column (Cytiva) with 30-35 ml ceramic hydroxyapatite (Bio-Rad, cat no. 1582000). PlyCB_{WT} was eluted from the column with three column volumes of 1 M sodium phosphate buffer (pH 7.2). Protein was subsequently dialyzed in PBS, 10% glycerol and stored at -80°C until use. Protein purification of PlyCB_{R66E} was performed as PlyCB_{WT} [13]. The fluorescent labeling of PlyCB_{WT} and PlyCB_{R66E} was performed using the manufacturer's recommended guidelines (Thermo Fisher, cat. No. A20174). Analytical gel filtration was used to determine the multimeric nature of PlyCB and PlyCB_{R66E}. A total of 100 µl (0.1 mg) of each protein was applied to a Superose 12 10/300 GL column (Cytiva) and run in isocratic conditions in PBS buffer for 1.5 column volumes on an AKTA FPLC system (Cytiva). Gel filtration standards (Bio-Rad) were also run under identical conditions.

159

160 **Purification of GAC**

161 GAS cells were grown overnight in THY media at 37°C. Cultures were centrifuged at 4,500 x g.
162 Pellets were washed and resuspended in 40 ml distilled water per each original liter of media
163 used and combined in an 800 ml beaker. 22.5 ml 4N sodium nitrite (5 ml per liter of culture) was
164 added to the beaker in addition to 22.5 ml glacial acetic acid (5 ml per liter of culture). An orbital
165 shaker was used to gently mix the beaker for 15 min in a hood. The mixture was centrifuged in
166 500 ml bottles at 8000 x g for 15 min. The supernatant was decanted to a new beaker and
167 neutralized with 1M sodium hydroxide. The total solution, about 300 ml, was filtered with a 0.45-
168 micron filter assembly. 50-50 ml aliquots were deposited in a 3.5 kDa membrane and dialyzed in
169 a 4-liter beaker overnight with water. The following day, the solution was concentrated using an
170 Amicon 400 ml stirred cell (model 8400) filter assembly with a 76 mm diameter Ultracel® 5 kDa
171 ultrafiltration disc (Millipore Sigma, cat. No. PLCC07610) for 2 hours with 60 psi. The ~10 ml
172 concentrate was loaded onto an S-100 column for final purification. Fractions were assayed for
173 Rha, lyophilized and stored at 4°C until use.

174

175 **Calculation of pRha concentration**

176 A modified protocol by Edgar *et al.* [21] based on a protocol from DuBois *et al.* [29] was used to
177 determine Rha concentration in purified GAC. Briefly, anthrone reagent was prepared by
178 dissolving 0.2% w/w anthrone in H₂SO₄. Eighty microliters of aqueous samples or standards
179 containing either GAC or L-Rha at known concentrations were added to a 1.5 ml microfuge tube.
180 To this same tube, 320 µL of the anthrone reagent was added. Samples were boiled at 98°C for
181 10 min in a heat block. Samples were cooled to room temperature, transferred to a quartz plate,
182 and the absorbance at 580 nm was recorded using a spectrophotometer. Rha concentration was
183 interpolated using an L-Rha standard curve.

184

Precipitation of PlyCB with GAC

PlyCB_{WT} and PlyCB_{R66E} samples were defrosted from storage at -80°C. Lyophilized GAC was resuspended in PBS. Both proteins and GAC were added to a 3.5 kDa dialysis membrane and dialyzed overnight in PBS. Protein concentrations were determined using a NanoDrop spectrophotometer (Thermo Fisher ND-2000) at 280 nm and were diluted to 5 mg/ml. The GAC was also assayed and diluted with PBS to 1.6 mg/ml. One-hundred microliters of proteins and 100 µl of GAC or PBS were mixed in a 250 µl quartz plate and allowed to incubate without shaking at room temperature. Visible precipitate formed in samples in 5-8 minutes. After recording the precipitate at 340 nm using a Spectramax® M5 (Molecular Devices) spectrophotometer, the total sample volume was transferred to a 1.5 ml microfuge tube. Samples were centrifuged at 14,000 x g to pellet the precipitate and supernatants were transferred to new 1.5 ml microfuge tubes. Two-hundred microliters of 8 M urea was added to the pellet. Pellets were resuspended and 5 µl of either pellet or supernatant were added to 40 µl water with 8 µl 6x Laemmli buffer with DTT. Samples were boiled at 98°C for 8 min, and then 12.5 µl were loaded onto a 7.5% SDS-PAGE and run for 32 min at 200 V. Proteins were visualized using Coomassie stain.

Lysis assay

A turbidity reduction assay was used to ascertain strain sensitivity to PlyC. This assay was performed as previously described [30], except PlyC was used at 2 µM. Eight technical replicates were performed.

Sensitivity of streptococcal species to PlyC-mediated lysis was analyzed using a wide range of clinical isolates: 1) GAS isolates: M1_{UK}, WT 5448 strain [M1], delta*gacI* 5448 strain [*dgacI*]; 2) GCS isolate: stC74A.0; 3) SDSE_*gac* isolates: stG245.0, stG652.0, stG485.0; GGS: stG6.0, stG485.0 and 4) *S. mutans* serotype c were used as negative controls. Briefly, all streptococcal strains were grown in THY at 37°C overnight in 5% CO₂, except for *S. mutans*, which was grown

in THB media. Next day, the bacterial cells were inoculated in 1:100 fresh media and grown until mid-logarithmic phase (OD₆₀₀ 1.0). The cells were washed in PBS and resuspended to an OD₆₀₀ of 2.0. In a 96-well plate, to a 100 µl of bacterial cells, 100 µl of PlyC [1 µg/ml] was added and immediately read at an absorbance of OD₆₀₀. The obtained values were standardized by subtracting from the background values. The data is plotted using GraphPad Prism version 9.

SDS-PAGE and blotting analysis

PlyCB_{WT} binding to recombinant *E. coli* expressing pRha was conducted using blot analysis. Briefly, the lysate from the overnight cultures was analyzed in 20% tricine gels. SDS-PAGE and protein transfers were performed according to manufactures instructions, Atto Ae-6050 Mini Gel chamber and Novex protein separation from Thermo Fisher, respectively. The PVDF membranes were blocked with 5% non-fat dry milk with Tris-Buffered Saline, 0.1% Tween® 20 detergent prior to incubation with PlyCB_{WT} labelled with Alexa Fluor® 647 (1:1000) for one hour at room temperature. Goat anti-rabbit GAC antibodies conjugated with IRDye® 800CW were used as a positive control (abcam ab216773). The resulting blots were imaged using the Licor Odyssey FC Imaging System. All the blots were processed in parallel under the same conditions.

Microscopy

Microscopic analysis of *E. coli* bacteria was performed using cells that were grown overnight in LB supplemented with antibiotics at 37°C and diluted 1:100 the next day and regrown until OD₆₀₀ reached 0.5. The cells were washed twice with PBS for 5 min at 10,000 rpm and stained with 1:1000 dilution of PlyCB^{AF647} and left for 20 minutes on ice in the darkness. Prior to fixing the cells with 4% paraformaldehyde, the cells were washed again with PBS. The fixed cells were mounted on 1% agarose coated microscopic slides and viewed under the CY5 channel on a fluorescent Deltavision widefield microscope.

FACS analysis

E. coli cells were grown overnight as described above and diluted 1:100 the next day, grown at 37°C and used for the assay at OD₆₀₀ = 0.5. The cells were washed with PBS and probed with PlyCB^{AF647} or PlyCB_{R66E}^{AF647} at 1:1000 dilution at 1:1000 dilution. Anti-GAC antibodies conjugated with FITC (ABIN238144, antibodies-online, titer 1:50) were used as a positive control. The samples were incubated for 20 minutes on ice at dark conditions. The cells were washed twice with PBS at 14,000 rpm for 5 minutes and fixed with 4% paraformaldehyde. BD LSRFortessa Flow Cytometry software was used to analyse the samples and the data interpretation was conducted with FlowJo™ software v10.6.2.

PlyC hydrolysis of sacculi

S. mutans wild-type (WT) and the mutant strain sacculi were obtained by the SDS-boiling procedure [25] followed by four washes each with 1 M NaCl and distilled water. The sacculi were resuspended in PBS to OD₆₀₀ of 1.0 and incubated with PlyC (5 µg/ml) in a 96-well plate. The lysis was monitored after 10, 20, 30, 40, 50 and 60 min as a decrease in OD₆₀₀. Results were reported as fold change in OD₆₀₀ of the sacculi incubated with PlyC vs the sacculi incubated without PlyC.

Results & Discussion

Pathogenic streptococci producing the GAC are susceptible to PlyC

A major component of the GAS cell wall is the GAC, building approximately 50% of the cell wall by weight [31]. The GAC is universally conserved amongst all isolated GAS strains on the basis of the gene cluster sequence conservation [32]. A number of *S. dysgalactiae* subsp. *equisimilis* (SDSE) isolates, naturally belonging to Group G Streptococci (GGS), have been reported to have

undergone homologous recombination and replaced their Group G Carbohydrate (GGC) gene cluster in parts with the GAC gene cluster [33-35]. We therefore expanded the previously reported PlyC streptococci cell lysis assay used by Nelson *et al.* [2] to investigate those new isolates named SDSE_*gac*. We also tested if PlyC was able to lyse a selection of GAS serotypes including a newly emerged isolate M1_{UK} [36], and included negative controls GGS isolates and *S. mutans* serotype C (Fig. 1B). In agreement with the published literature, all tested GAS serotypes are susceptible to PlyC lysis and the two GGS isolates are resistant to PlyC. The GGC does not contain the GAC, GCC and SCC pRha backbone with $\rightarrow 3)\alpha$ -Rha(1 \rightarrow 2) α -Rha(1 \rightarrow di-saccharide repeats, but an alternating Rha-Gal carbohydrate [37] (Fig. 1A). All three SDSE_*gac* isolates tested have inherited parts of the *gac* gene cluster and produce the GAC instead of the GGC [34]. Strikingly, the SDSE_*gac* strains are all sensitive to PlyC treatment. The fact that PlyC is able to lyse SDSE_*gac* strains that express the GAC, but PlyC does not lyse the isogenic GGS strains producing the GGC strongly suggests that the GAC is a critical component of PlyC recognition and subsequent activity.

Importantly, all strains tested in this study that are susceptible to PlyC lysis produce a cell wall polysaccharide that contains the pRha backbone and a β -linked sugar substituent on the α 1,2-linked Rha (Fig 1A). We therefore suggest that the pRha backbone with and without a side chain are both vital ligands to assist PlyC activity and the new SDSE_*gac* isolates will also be susceptible to PlyC treatment due to production of the GAC.

Purified GAC precipitates PlyCB - but not PlyCB_{R66E}

The lysis assay of GAS cells, and in particular of the SDSE_*gac* variants, suggests that either the ubiquitous pRha or GAC in GAS cells is the ligand for PlyCB. We propose that the PlyCB octameric CBD binds GAC and/or the GAC pRha backbone. We tested this hypothesis by investigating the binding of PlyC to partially purified GAC. We hypothesized if the GAC was able

to precipitate PlyCB, an interaction of the two systems must have occurred [38]. As a negative control, we employed the previously published inactive mutant PlyCB_{R66E}, which lost the ability to bind to GAS cells [13]. Consistent with prior analytical gel filtration and dynamic light scattering results [12], as well as the crystal structures [13, 39, 40], both PlyCB and its R66E mutant self-assemble into stable octameric structures (Fig 1C, SF 1). The purified proteins were incubated with the extracted GAC, and precipitation was monitored at 340 nm, a standard wavelength for measuring protein aggregation [41, 42] (Fig. 2A, B). Whilst keeping the PlyCB concentration constant, we varied the concentration of GAC. Within five minutes at room temperature the solution became turbid, suggesting aggregation (Fig. 2A). When the PlyCB concentration was kept constant and the GAC concentration was varied, the turbidity correlated with PlyCB concentration in a dose dependent manner, suggesting that PlyCB requires GAC to aggregate. Importantly, PlyCB did not self-aggregate when no GAC was added in the assay. Furthermore, no aggregation was detected when PlyCB_{R66E} was incubated with purified GAC (Fig. 2B). To demonstrate the presence of PlyCB in the precipitates, we analyzed the soluble and pellet fractions (Fig. 2C, D). A higher yield of aggregated PlyCB was found in the pelleted samples when compared to the soluble fraction (Fig. 2D). A similar precipitation effect was observed when we varied the PlyCB concentration and kept the GAC concentration constant (Fig. 2E, F), demonstrating that both species are necessary for an interaction.

PlyCB binds to recombinantly produced pRha backbone

The purified GAC from bacteria contains a mixture of carbohydrates, including the fully decorated GAC with GroP[21] and a small proportion of the polysaccharide lacking the side chains [43]. PlyCB is able to lyse a number of GAS mutants including GAVS and *dgacI*_M1 [2, 11, 25, 44], that decorate the cell wall with the unmodified GAC lacking the side chains (Fig. 1A, B), suggesting that the pRha backbone of the GAC is the minimal carbohydrate structure required for PlyCB

binding. To test this hypothesis, we recombinantly produced the pRha backbone in *E. coli* cells. We and others have previously reported that the *S. mutans* *sccABCDEFGF* gene cluster, when transformed into *E. coli* cells, functionally produces the pRha backbone attached to the lipid A [23, 28]. Additionally, to understand if PlyCB recognizes a specific pRha backbone, we engineered *E. coli* cells expressing the GAC gene cluster *gacABCDEFGF* required for the GAC pRha production. *E. coli* cells carrying an empty plasmid were used as a negative control.

Next, we investigated the binding of PlyCB conjugated with Alexa Fluor® 647 (PlyCB^{AF647}) to an *E. coli* total cell lysate expressing the pRha backbone of the SCC or GAC, respectively (Fig. 3A). The blotted membranes were incubated with PlyCB^{AF647}, and a positive interaction between PlyCB^{AF647} and the *E. coli* produced pRha is visualized as a prominent band around 40 kDa. The size of the band agrees with the band detected by anti-GAC antibodies that were previously reported to recognize pRha [23] (Fig. 3B). Importantly, PlyCB^{AF647} and GAC antibodies do not interact with the cell lysate of *E. coli* expressing an empty plasmid (Fig. 3 A, B). Contrary, the R66E-mutant protein was not able to detect the pRha of the PAGE separated sample (SF 2). We further confirmed the ability of PlyCB to bind to *E. coli* cells decorated with the pRha by fluorescent microscopy (SF 3). Only cells that produce the pRha are detected by the PlyCB^{AF647}, in agreement with the results of the blot analysis.

To gather additional evidence that PlyCB interacts with the pRha backbone, we established a flow cytometry assay to analyse the binding of PlyCB^{AF647} to pRha-producing *E. coli*. Unstained *E. coli* cells that express pRha or carry an empty plasmid sort in the identical range (Fig. 4A). The GAC antibodies label exclusively the cells producing pRha (Fig. 4A). A similar pattern of the GAC antibodies binding is observed when the cells were incubated with PlyCB^{AF647} (Fig. 4B). Contrary, the PlyCB_{R66E}^{AF647} mutant protein does not bind to *E. coli* cells, and PlyCB^{AF647} does not interact with the cells expressing an empty vector (Fig. 4B). Taken together, these data provide the first

definitive evidence that the pRha backbone of GAC and SCC is a binding receptor of the PlyCB octameric subunit.

PlyC lyses engineered *S. mutans* producing the GAC

Despite the fact that the SCC pRha backbone is identical to the GAC, *S. mutans* is resistant to PlyC lysis (Fig. 1A). To get a better understanding why *S. mutans* is resistant to PlyC, we compared PlyC-induced lysis of the sacculi purified from *S. mutans* WT and a number of mutant strains producing different SCC variants (Fig. 5). First, we examined the $\Delta sccH$ mutant producing the GroP-deficient SCC [21]. Similar to *S. mutans* WT, $\Delta sccH$ was resistant to PlyC-mediated lysis (Fig. 5). Second, we tested the *sccN* deletion mutant, $\Delta sccN$, that is deficient in the enzyme required for generation of the Glc side chains [45]. A time dependent lysis is observed for $\Delta sccN$. Expression of the WT copy of *sccN* in $\Delta sccN$ (the $\Delta sccN$:*psccN* strain) fully restored the resistance of the bacteria to PlyC (Fig. 5). These observations clearly suggest that PlyC is able to bind to the *S. mutans* cells producing the unmodified pRha backbone, and the Glc side chains in SCC hinders PlyC binding. We then investigated whether the addition of the GAC GlcNAc side chains to the pRha backbone affects sensitivity of the engineered *S. mutans* sacculi to PlyC-induced lysis. We expressed the GAS genes *gacHIJKL* required for the formation and addition of the GlcNAc side chains and GroP to GAC [21], in the $\Delta sccN$ background strain in two versions, creating the $\Delta sccN$:*pgacHI*JKL* and $\Delta sccN$:*pgacHIJKL* strains [45]. The plasmid *pgacHI*JKL* contains an inserted stop codon in the *gacI* gene required for generation of the GlcNAc side chain, and, therefore, the $\Delta sccN$:*pgacHI*JKL* strain produces the unmodified SCC lacking any side chains (Fig. 5). As expected, the sacculi isolated from this strain remains susceptible to PlyC lysis. We previously showed that in $\Delta sccN$:*pgacHIJKL*, the Glc side chains are replaced with the GlcNAc side chains [45]. Interestingly, expression of *gacHIJKL* in $\Delta sccN$ did not restore the resistance of the bacteria to PlyC (Fig. 5), indicating that the GlcNAc side chains do not obstruct PlyC binding. Lastly, we analyzed PlyC-mediated lysis of sacculi purified from the $\Delta rgpG$ mutant, which is

deficient in SCC expression [45]. The RgpG protein catalyzes the first step in SCC biosynthesis [46]. In comparison to $\Delta sccN$, PlyC-induced lysis of $\Delta rgpG$ was less pronounced (Fig. 5), indicating the importance of the pRha backbone of SCC in PlyC activity and supporting the findings that the pRha backbone is a ligand contributing to PlyC binding. These studies reveal that if the SCC is 'unmasked' *i.e.*, stripped of the Glc and Glc-GroP side chains, it becomes a ligand for PlyCB and that *S. mutans* is PlyC susceptible if SCC is replaced with GAC.

Concluding remarks

A structural feature of the PlyCB protein remains to be discovered that explains why only certain streptococci are susceptible to PlyC's lytic activity. The pRha decorated with an α -linked side chain sugar appears not compatible with the PlyCB ligand binding site and therefore only those streptococci expressing pRha decorated with β -linked substituents, such as GlcNAc and GlcNAc-GroP are susceptible. This could potentially be exploited for diagnostic purposes or in the case of SCC and Group G Streptococci, opens up the potential for novel therapeutic approaches. If SCC was treated by PlyC in combination with an additional enzyme that removes the α -linked side chains in these streptococcal carbohydrates, this would expose the pRha backbone and subsequently make these strains susceptible. A recently accepted manuscript by Boendum *et al.* [40] reported 19 potential binding states of tetra-rhamnose to PlyCB. Once the pRha ligand site has been identified, this could also be further exploited by directed evolution approaches to generate PlyCB protein variants that are capable to bind the carbohydrates from, for example, SCC and Group G Streptococci.

Whilst much has been learned about the structure and function of PlyC in the past 20 years, many questions remain, specifically with respect to its interaction with the PG. Considering the average

length of the cellular pRha is 7-10 kDa [25] and that the α 1,2-1,3-pRha with or without β -configured GlcNAc/GalNAc-side chains bind to PlyCB, it is inviting to speculate that an element of avidity is responsible for tight binding of the PlyCB octamer to the streptococcal surface, in agreement with the recent accepted manuscript [40]. Further proof is needed to substantiate this hypothesis. Another question lies in the actions of the EADs relative to the PlyCB octamer. PlyC clearly has a high turnover as demonstrated in multiple biochemical assays. However, it is unknown if PlyCB “docks” to the surface and the flexibility of the holoenzyme allows the PlyC EADs to cleave multiple bonds in a localized area weakening the overall superstructure of the PG. Alternatively, the enzymatic turnover could be dictated by a balance of on and off rates of the EADs and CBD monomers leading to widespread hydrolysis of the PG. Lastly, it is unknown whether PlyC binds, cleaves, and releases the PG at random points on the streptococcal surface or works its way down a single strand of PG in a processive manner. It is noteworthy that cellulase enzymes, which cleave the β 1,4 glycosidic linkages in cellulose, possess a catalytic domain, a flexible linker, and a cellulose binding domain, analogous to the traditional endolysins. It has been demonstrated that energy is stored in the flexible linker can adopt compact and extended configurations that allows the cellulase to move in a “caterpillar-like” motion down a chain of cellulose [47, 48]. Although PlyC does not contain an equivalent flexible linker, the octameric nature of PlyCB invites the possibility that it may interact with the successive pRha strands allowing PlyC to depolymerize the PG in a zipper-like fashion.

In conclusion, the α 1,2-1,3-pRha is the definite, minimal carbohydrate substrate for the PlyCB subunit. We validated this by comprehensive experiments, using pRha recombinantly produced in *E. coli*, and the *S. mutans* variants of the Rha-based polysaccharides. The work described here provides valuable insight into the molecular interactions that define a PlyC’s host specificity, which can inform the future studies as well as engineering approaches.

418

419 **Data Availability Statement**

420 No mandated datasets are associated with the paper.

421

422

423

424

425

426

427

428

429 **Figure Legend**

430 Fig 1)

431 A) Symbolic drawings of the carbohydrate structures of GAS (GAC), polyrhamnose backbone
432 (pRha) in GAVS and M1_*dgacI*, GCS (GCC), GGS (GGC) and *S. mutans* (SCC). For simplicity,
433 the reported glycerol phosphate of occasionally present on the GAC and SCC side chains have
434 been omitted. Repeat units are marked with brackets. The pRha backbone with alternating (α 1-
435 \rightarrow 2) and (α 1- \rightarrow 3) linkages, whilst the α 1- \rightarrow 2 Rha being decorated with a β 1- \rightarrow 3 side chain is a
436 commonality among all PlyC susceptible strains. B) PlyC lysis of streptococcal pathogens. Group
437 A and Group C Streptococci serotypes are susceptible to PlyC-mediated lysis. Group G
438 Streptococci show limited susceptibility and *S. mutans* is resistant to PlyC lysis. Three SDSE
439 isolates that produce the GAC instead of GGC (SDSE_*gac*) are susceptible to PlyC treatment. C)
440 Bottom and top view onto WT PlyCB octameric structure with the GlyH and EAD domains omitted
441 (4f78.pdb).

442

443 Fig 2)

Precipitation studies of purified PlyCB and GAC reveal direct interaction of PlyCB with GAC. A) The PlyCB concentration is kept constant whilst the GAC concentration is varied. Visible precipitate forms at the higher concentrations. B) The precipitate level is measured spectrophotometrically at 340 nm and compared to the mutant PlyCB_{R66E}, which does not bind the GAC. C) Coomassie stained and D) densitometry analysis of PlyCB protein from the supernatant fraction and aggregates (pellets). E, F) The same dose dependency is observed when the PlyCB concentration is varied. Arrowhead depicts PlyCB protein at 8 kDa.

Fig 3)

Representation of immunoblot analysis of the cell lysate of *E. coli* expressing the SCC and GAC pRha and carrying an empty control plasmid (-ve). A) Blot was incubated with PlyCB-Alexa Fluor® 647 (PlyCB_{WT}^{AF647}). B) Probing the same samples with the GAC antibodies confirms the presence of GAC in the bands. Molecular mass markers are given in kDa.

Fig 4)

PlyCB binding to *E. coli* cells were investigated by flow cytometry after labelling with PlyCB_{WT}^{AF647} and PlyCB_{R66E}^{AF647} mutant proteins. Blue: -ve control cells without pRha. Red: pRha producing *E. coli* cells. Representative histograms are shown. A) Left panel: unstained cells. Right panel: The anti-GAC antibodies (GAC-FITC) were used as a positive control to label the *E. coli* cells producing pRha. The antibodies do not bind to the *E. coli* cells carrying an empty plasmid (-ve). B) Left panel: PlyCB_{WT}^{AF647} binds to the *E. coli* cells producing pRha, but not to the *E. coli* cells carrying an empty plasmid (-ve). Right panel: PlyCB_{R66E}^{AF647} does not binds to the *E. coli* cells producing pRha.

Fig 5)

The PlyC-mediated lysis of sacculi purified from *S. mutans* strains. The lysis was monitored after 10, 20, 30, 40, 50 and 60 min as a decrease in OD₆₀₀. Results are presented as a fold change in OD₆₀₀ of the sacculi incubated with PlyC vs. the sacculi incubated without PlyC. Data points and error bars represent mean values of four biological replicates and standard deviation, respectively. P-values were determined by two-way ANOVA with Dunnett's multiple comparisons test.

FUNDING

The HCD laboratory is supported by Wellcome and Royal Society Grant 109357/Z/15/Z and the University of Dundee Wellcome Trust Funds 105606/Z/14/Z and Tenovus Scotland Large Research Grant [T17/17]. HK is supported by the National Institute of Standards and Technology (NIST). VAF is supported by Rockefeller University laboratory funds. NK is supported by NIH grants R01 DE028916 from the NIDCR and R01 AI143690 from the NIAID.

ACKNOWLEDGEMENTS

We thank the CDC Streptococcus Lab and the Active Bacterial Core surveillance program for sharing isolates isolate 20170556 (stG6.0), 20171682 (stC74A.0), 20154376 (stG245.0), 20173686 (stG652.0), 20170560 and 20176966 (stG485.0). We also thank Ryan Heselpoth, Yang Shen, and Sara Linden for reagents, technical advice, and helpful discussion. For the purpose of open access, the authors have applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

References

494 1 Fernandes, S. and Sao-Jose, C. (2018) Enzymes and Mechanisms Employed by Tailed
 495 Bacteriophages to Breach the Bacterial Cell Barriers. *Viruses*. **10**
 496 2 Nelson, D., Loomis, L. and Fischetti, V. A. (2001) Prevention and elimination of upper
 497 respiratory colonization of mice by group A streptococci by using a bacteriophage lytic enzyme.
 498 *Proc Natl Acad Sci U S A*. **98**, 4107-4112
 499 3 Dams, D. and Briers, Y. (2019) Enzybiotics: Enzyme-Based Antibacterials as
 500 Therapeutics. *Adv Exp Med Biol*. **1148**, 233-253
 501 4 Love, M. J., Bhandari, D., Dobson, R. C. J. and Billington, C. (2018) Potential for
 502 Bacteriophage Endolysins to Supplement or Replace Antibiotics in Food Production and Clinical
 503 Care. *Antibiotics (Basel)*. **7**
 504 5 Schmelcher, M. and Loessner, M. J. (2020) Bacteriophage endolysins - extending their
 505 application to tissues and the bloodstream. *Curr Opin Biotechnol*. **68**, 51-59
 506 6 Heselpoth, R. D., Swift, S. M., Linden, S. B., Mitchell, M. S. and Nelson, D. C. (2018)
 507 Enzybiotics: endolysins and bacteriocins. In *Bacteriophages - biology, technology, and therapy*
 508 (Harper, D., Abedon, S. T., Burrowes, B. and McConvill, M., eds.), Springer International
 509 Publishing, Dordrecht, The Netherlands
 510 7 Loessner, M. J., Kramer, K., Ebel, F. and Scherer, S. (2002) C-terminal domains of
 511 *Listeria monocytogenes* bacteriophage murein hydrolases determine specific recognition and
 512 high-affinity binding to bacterial cell wall carbohydrates. *Molecular microbiology*. **44**, 335-349
 513 8 Mesnage, S., Dellarole, M., Baxter, N. J., Rouget, J. B., Dimitrov, J. D., Wang, N.,
 514 Fujimoto, Y., Hounslow, A. M., Lacroix-Desmazes, S., Fukase, K., Foster, S. J. and Williamson,
 515 M. P. (2014) Molecular basis for bacterial peptidoglycan recognition by LysM domains. *Nat*
 516 *Commun*. **5**, 4269
 517 9 Evans, A. C. (1934) *Streptococcus* bacteriophage: a study of four serological types.
 518 *Public Health Rep*. **49**, 1386-1401
 519 10 Krause, R. M. (1957) Studies on bacteriophages of hemolytic streptococci. I. Factors
 520 influencing the interaction of phage and susceptible host cell. *J Exp Med*. **106**, 365-384
 521 11 Maxted, W. R. (1957) The active agent in nascent phage lysis of streptococci. *J Gen*
 522 *Microbiol*. **16**, 584-595
 523 12 Nelson, D., Schuch, R., Chahales, P., Zhu, S. and Fischetti, V. A. (2006) PlyC: a
 524 multimeric bacteriophage lysin. *Proc Natl Acad Sci U S A*. **103**, 10765-10770
 525 13 McGowan, S., Buckle, A. M., Mitchell, M. S., Hoopes, J. T., Gallagher, D. T., Heselpoth,
 526 R. D., Shen, Y., Reboul, C. F., Law, R. H., Fischetti, V. A., Whisstock, J. C. and Nelson, D. C.
 527 (2012) X-ray crystal structure of the streptococcal specific phage lysin PlyC. *Proc Natl Acad Sci*
 528 *U S A*. **109**, 12752-12757
 529 14 Lancefield, R. C. (1933) A Serological Differentiation of Human and Other Groups of
 530 Hemolytic Streptococci. *J Exp Med*. **57**, 571-595
 531 15 Vandamme, P., Pot, B., Falsen, E., Kersters, K. and Devriese, L. A. (1996) Taxonomic
 532 study of lancefield streptococcal groups C, G, and L (*Streptococcus dysgalactiae*) and proposal
 533 of *S. dysgalactiae* subsp. *equisimilis* subsp. nov. *Int J Syst Bacteriol*. **46**, 774-781
 534 16 Coligan, J. E., Kindt, T. J. and Krause, R. M. (1978) Structure of the streptococcal
 535 groups A, A-variant and C carbohydrates. *Immunochemistry*. **15**, 755-760
 536 17 Pritchard D G, G. R. L., Michlek S M, McGhee J R. . (1986) Biochemical aspects of
 537 serotype carbohydrate antigens of *Streptococcus mutans*. In: Hamada S, Michalek S M, Kiyono
 538 H, Menaker L, McGhee J R, editors. *Molecular microbiology and immunology of Streptococcus*
 539 *mutans*. Amsterdam, The Netherlands: . Elsevier Science Publishers; , pp. 39-49.
 540 18 Coligan, J. E., Schnute, W. C., Jr. and Kindt, T. J. (1975) Immunochemical and chemical
 541 studies on streptococcal group-specific carbohydrates. *J Immunol*. **114**, 1654-1658
 542 19 Huang, D. H., Rama Krishna, N. and Pritchard, D. G. (1986) Characterization of the
 543 group A streptococcal polysaccharide by two-dimensional ¹H-nuclear-magnetic-resonance
 544 spectroscopy. *Carbohydrate research*. **155**, 193-199

545 20 Neiwert, O., Holst, O. and Duda, K. A. (2014) Structural investigation of rhamnose-rich
546 polysaccharides from *Streptococcus dysgalactiae* bovine mastitis isolate. *Carbohydr Res.* **389**,
547 192-195

548 21 Edgar, R. J., van Hensbergen, V. P., Ruda, A., Turner, A. G., Deng, P., Le Breton, Y.,
549 El-Sayed, N. M., Belew, A. T., McIver, K. S., McEwan, A. G., Morris, A. J., Lambeau, G.,
550 Walker, M. J., Rush, J. S., Korotkov, K. V., Widmalm, G., van Sorge, N. M. and Korotkova, N.
551 (2019) Discovery of glycerol phosphate modification on streptococcal rhamnose
552 polysaccharides. *Nature chemical biology.* **15**, 463-471

553 22 van der Beek, S. L., Zorzoli, A., Canak, E., Chapman, R. N., Lucas, K., Meyer, B. H.,
554 Evangelopoulos, D., de Carvalho, L. P. S., Boons, G. J., Dorfmüller, H. C. and van Sorge, N.
555 M. (2019) Streptococcal dTDP-L-rhamnose biosynthesis enzymes: functional characterization
556 and lead compound identification. *Molecular microbiology.* **111**, 951-964

557 23 Zorzoli, A., Meyer, B. H., Adair, E., Torgov, V. I., Veselovsky, V. V., Danilov, L. L., Uhrin,
558 D. and Dorfmüller, H. C. (2019) Group A, B, C, and G *Streptococcus* Lancefield antigen
559 biosynthesis is initiated by a conserved alpha-d-GlcNAc-beta-1,4-l-rhamnosyltransferase. *J Biol*
560 *Chem.* **294**, 15237-15256

561 24 van der Beek, S. L., Le Breton, Y., Ferenbach, A. T., Chapman, R. N., van Aalten, D. M.,
562 Navratilova, I., Boons, G. J., McIver, K. S., van Sorge, N. M. and Dorfmüller, H. C. (2015)
563 GacA is essential for Group A *Streptococcus* and defines a new class of monomeric dTDP-4-
564 dehydrorhamnose reductases (RmlD). *Molecular microbiology.* **98**, 946-962

565 25 Rush, J. S., Edgar, R. J., Deng, P., Chen, J., Zhu, H., van Sorge, N. M., Morris, A. J.,
566 Korotkov, K. V. and Korotkova, N. (2017) The molecular mechanism of N-acetylglucosamine
567 side-chain attachment to the Lancefield group A carbohydrate in *Streptococcus pyogenes*. *J*
568 *Biol Chem.* **292**, 19441-19457

569 26 Lood, R., Raz, A., Molina, H., Euler, C. W. and Fischetti, V. A. (2014) A highly active and
570 negatively charged *Streptococcus pyogenes* lysin with a rare D-alanyl-L-alanine endopeptidase
571 activity protects mice against streptococcal bacteremia. *Antimicrob Agents Chemother.* **58**,
572 3073-3084

573 27 Gera, K. and McIver, K. S. (2013) Laboratory growth and maintenance of *Streptococcus*
574 *pyogenes* (the Group A *Streptococcus*, GAS). *Curr Protoc Microbiol.* **30**, 9D 2 1-9D 2 13

575 28 Shibata, Y., Yamashita, Y., Ozaki, K., Nakano, Y. and Koga, T. (2002) Expression and
576 characterization of streptococcal *rgp* genes required for rhamnan synthesis in *Escherichia coli*.
577 *Infect Immun.* **70**, 2891-2898

578 29 DuBois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A. and Smith, F. (1956) Colorimetric
579 method for determination of sugars and related substances. *Anal. Chem.* **28**, 350-356

580 30 Heselpoth, R. D., Yin, Y., Moulton, J. and Nelson, D. C. (2015) Increasing the stability of
581 the bacteriophage endolysin PlyC using rationale-based FoldX computational modeling. *Protein*
582 *Eng Des Sel.* **28**, 85-92

583 31 M., M. (1952) The lysis of group A hemolytic streptococci by extracellular enzymes of
584 *Streptomyces albus*. II. Nature of the cellular substrate attacked by the lytic enzymes. . *J Exp*
585 *Med.* . **96**, 569-580

586 32 Davies, M. R., McIntyre, L., Mutreja, A., Lacey, J. A., Lees, J. A., Towers, R. J.,
587 Duchene, S., Smeesters, P. R., Frost, H. R., Price, D. J., Holden, M. T. G., David, S., Giffard, P.
588 M., Worthing, K. A., Seale, A. C., Berkley, J. A., Harris, S. R., Rivera-Hernandez, T., Berking,
589 O., Cork, A. J., Torres, R., Lithgow, T., Strugnelli, R. A., Bergmann, R., Nitsche-Schmitz, P.,
590 Chhatwal, G. S., Bentley, S. D., Fraser, J. D., Moreland, N. J., Carapetis, J. R., Steer, A. C.,
591 Parkhill, J., Saul, A., Williamson, D. A., Currie, B. J., Tong, S. Y. C., Dougan, G. and Walker, M.
592 J. (2019) Author Correction: Atlas of group A streptococcal vaccine candidates compiled using
593 large-scale comparative genomics. *Nat Genet.* **51**, 1295

- 33 Brandt, C. M., Haase, G., Schnitzler, N., Zbinden, R. and Luttkicken, R. (1999) Characterization of blood culture isolates of *Streptococcus dysgalactiae* subsp. *equisimilis* possessing Lancefield's group A antigen. *J Clin Microbiol.* **37**, 4194-4197
- 34 Chochua, S., Metcalf, B. J., Li, Z., Rivers, J., Mathis, S., Jackson, D., Gertz, R. E., Jr., Srinivasan, V., Lynfield, R., Van Beneden, C., McGee, L. and Beall, B. (2017) Population and Whole Genome Sequence Based Characterization of Invasive Group A *Streptococci* Recovered in the United States during 2015. *mBio.* **8**
- 35 McMillan, D. J., Vu, T., Bramhachari, P. V., Kaul, S. Y., Bouvet, A., Shaila, M. S., Karmarkar, M. G. and Sriprakash, K. S. (2010) Molecular markers for discriminating *Streptococcus pyogenes* and *S. dysgalactiae* subspecies *equisimilis*. *Eur J Clin Microbiol Infect Dis.* **29**, 585-589
- 36 Lynskey, N. N., Jauneikaite, E., Li, H. K., Zhi, X., Turner, C. E., Mosavie, M., Pearson, M., Asai, M., Lobkowicz, L., Chow, J. Y., Parkhill, J., Lamagni, T., Chalker, V. J. and Srisakandan, S. (2019) Emergence of dominant toxigenic M1T1 *Streptococcus pyogenes* clone during increased scarlet fever activity in England: a population-based molecular epidemiological study. *Lancet Infect Dis.* **19**, 1209-1218
- 37 Pritchard, D. G., Rener, B. P., Furner, R. L., Huang, D. H. and Krishna, N. R. (1988) Structure of the group G streptococcal polysaccharide. *Carbohydr Res.* **173**, 255-262
- 38 Mirica, K. A., Lockett, M. R., Snyder, P. W., Shapiro, N. D., Mack, E. T., Nam, S. and Whitesides, G. M. (2012) Selective precipitation and purification of monovalent proteins using oligovalent ligands and ammonium sulfate. *Bioconjug Chem.* **23**, 293-299
- 39 Shen, Y., Barros, M., Vennemann, T., Gallagher, D. T., Yin, Y., Linden, S. B., Heselpoth, R. D., Spencer, D. J., Donovan, D. M., Moul, J., Fischetti, V. A., Heinrich, F., Losche, M. and Nelson, D. C. (2016) A bacteriophage endolysin that eliminates intracellular streptococci. *Elife.* **5**
- 40 Broendum, S. S., Williams, D. E., Hayes, B. K., Kraus, F., Fodor, J., Clifton, B. E., Geert Volbeda, A., Codee, J. D. C., Riley, B. T., Drinkwater, N., Farrow, K. A., Tsyganov, K., Heselpoth, R. D., Nelson, D. C., Jackson, C. J., Buckle, A. M. and McGowan, S. (2021) High avidity drives the interaction between the streptococcal C1 phage endolysin, PlyC, with the cell surface carbohydrates of Group A *Streptococcus*. *Molecular microbiology*
- 41 Sato, T., Mattison, K. W., Dubin, P. L., Kamachi, M. and Morishima, Y. (1998) Effect of protein aggregation on the binding of lysozyme to pyrene-labeled polyanions. *Langmuir.* **14**, 5430-5437
- 42 Wilcken, R., Wang, G., Boeckler, F. M. and Fersht, A. R. (2012) Kinetic mechanism of p53 oncogenic mutant aggregation and its inhibition. *Proc Natl Acad Sci U S A.* **109**, 13584-13589
- 43 Michon, F., Moore, S. L., Kim, J., Blake, M. S., Auzanneau, F. I., Johnston, B. D., Johnson, M. A. and Pinto, B. M. (2005) Doubly branched hexasaccharide epitope on the cell wall polysaccharide of group A streptococci recognized by human and rabbit antisera. *Infect Immun.* **73**, 6383-6389
- 44 van Sorge, N. M., Cole, J. N., Kuipers, K., Henningham, A., Aziz, R. K., Kasirer-Friede, A., Lin, L., Berends, E. T. M., Davies, M. R., Dougan, G., Zhang, F., Dahesh, S., Shaw, L., Gin, J., Cunningham, M., Merriman, J. A., Hutter, J., Lepenies, B., Rooijackers, S. H. M., Malley, R., Walker, M. J., Shattil, S. J., Schlievert, P. M., Choudhury, B. and Nizet, V. (2014) The classical lancefield antigen of group A *Streptococcus* is a virulence determinant with implications for vaccine design. *Cell Host Microbe.* **15**, 729-740
- 45 Svetlana Zamakhaeva, C. T. C., Jeffrey S. Rush, Sowmya Ajay Castro, Alexander E. Yarawsky, Andrew B. Herr, Nina M. van Sorge, Helge C. Dorfmueller, Gregory I. Frolenkov, Konstantin V. Korotkov, Natalia Korotkova. (2020) Modification of cell wall polysaccharide spatially controls cell division in *Streptococcus mutans*. *bioRxiv*
- 46 De, A., Liao, S., Bitoun, J. P., Roth, R., Beatty, W. L., Wu, H. and Wen, Z. T. (2017) Deficiency of RgpG causes major defects in cell division and biofilm formation, and deficiency of

645 LytR-CpsA-Psr family proteins leads to accumulation of cell wall antigens in culture medium by
 646 *Streptococcus mutans*. Appl Environ Microbiol. **83**
 647 47 Rabinovich, M. L., Melnick, M. S. and Bolobova, A. V. (2002) The structure and
 648 mechanism of action of cellulolytic enzymes. Biochemistry (Mosc). **67**, 850-871
 649 48 McCabe, C., Zhao, X., Adney, W. S., Himmel, M. E. (2010) Energy Storage in Cellulase
 650 Linker Peptides? Computational Modeling in Lignocellulosic Biofuel Production. **1052**, 119-134
 651

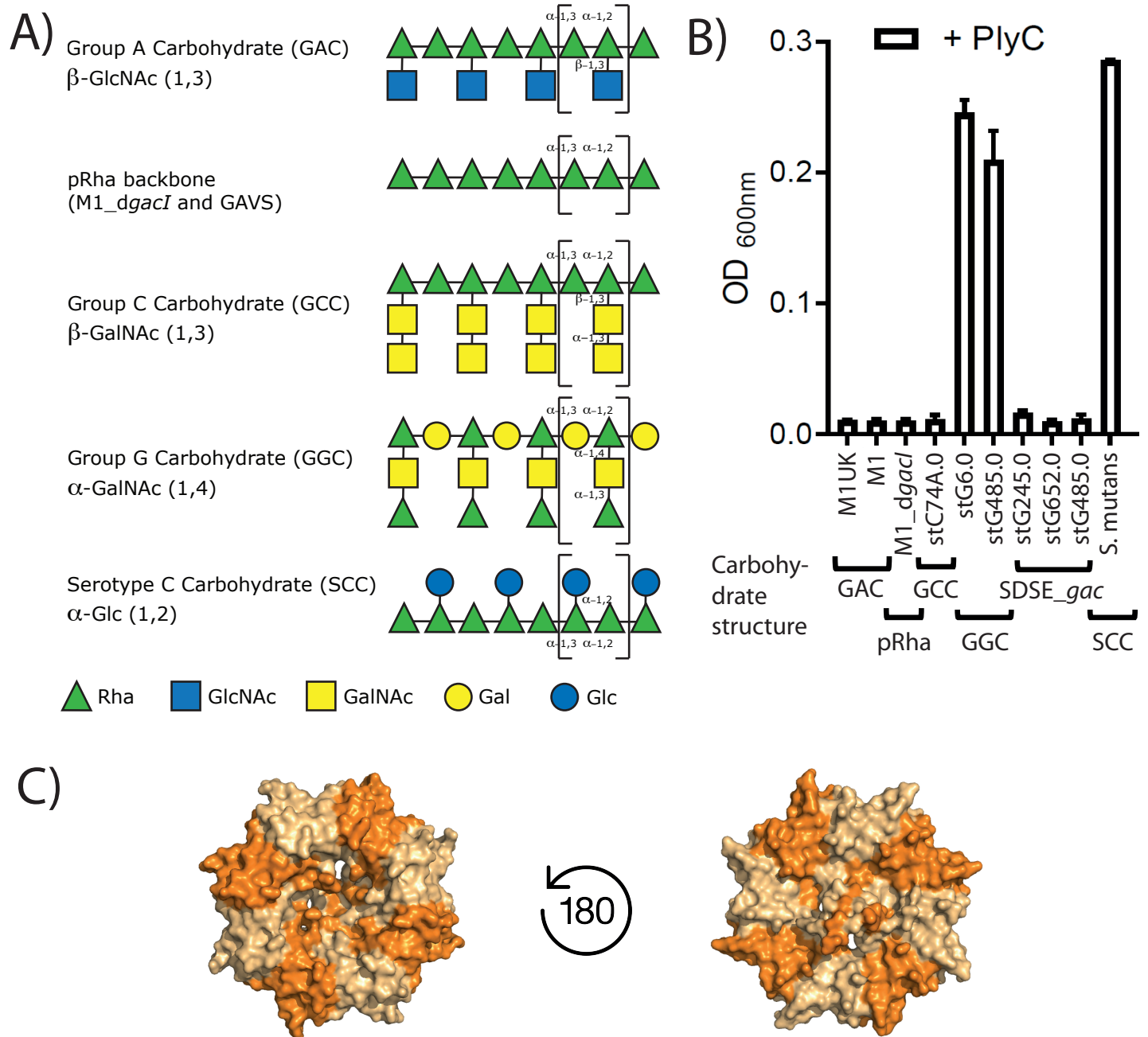


Fig 1)

A) Symbolic drawings of the carbohydrate structures of GAS (GAC), polyrhamnose backbone (pRha) in GAVS and M1_dgacI, GCS (GCC), GGS (GGC) and S. mutans (SCC). For simplicity, the reported glycerol phosphate of occasionally present on the GAC and SCC side chains have been omitted. Repeat units are marked with brackets. The pRha backbone with alternating (α 1 \rightarrow 2) and (α 1 \rightarrow 3) linkages, whilst the α 1 \rightarrow 2 rhamnose being decorated with a β 1 \rightarrow 3 side chain is a commonality among all PlyC susceptible strains. B) PlyC lysis of streptococcal pathogens. Group A and Group C Streptococci serotypes are susceptible to PlyC-mediated lysis. Group G Streptococci show limited susceptibility and S. mutans is resistant to PlyC lysis. Three SDSE isolates that produce the GAC instead of GGC (SDSE_gac) are susceptible to PlyC treatment. C) Bottom and top view of the WT PlyCB octameric ring structure with the GlyH and EAD domains omitted (4f78.pdb).

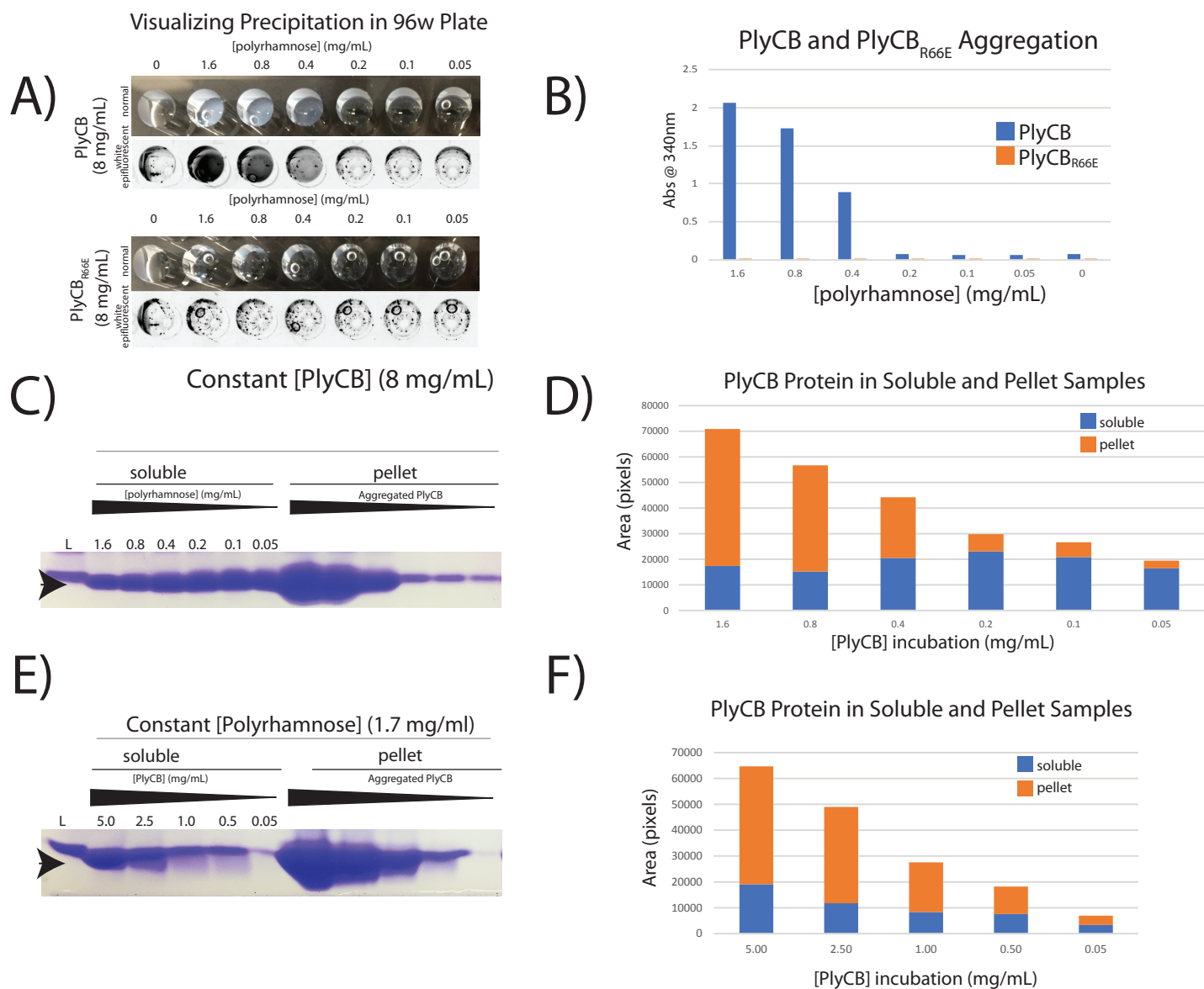


Fig 2) Precipitation studies of purified PlyCB and GAC reveal direct interaction of PlyCB with GAC. A) The PlyCB concentration is kept constant whilst the GAC concentration is varied. Visible precipitate forms at the higher concentrations. B) The precipitate level is measured spectrophotometrically at 340 nm and compared to the mutant PlyCB_{R66E}, which does not bind the GAC. C) Coomassie stained and D) densitometry analysis of PlyCB protein from the supernatant fraction and aggregates (pellets). E, F) The same dose dependency is observed when the PlyCB concentration is varied. Arrowhead depicts PlyCB protein (8 kDa).

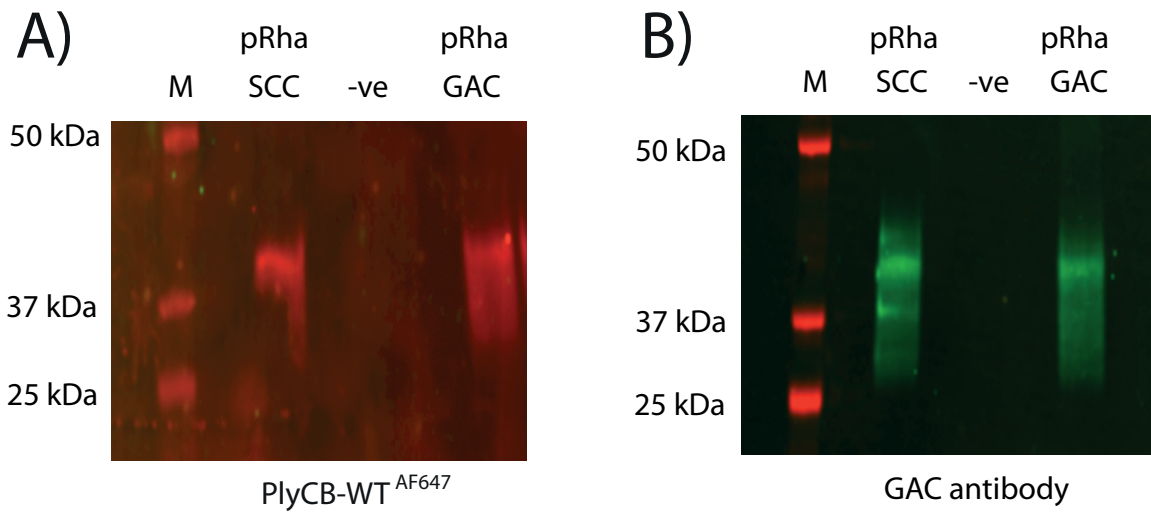


Fig 3)

Representation of immunoblot analysis of the cell lysate of *E. coli* expressing the SCC and GAC pRha and carrying an empty control plasmid (-ve). A) Blot was incubated with PlyCB-Alexa Fluor® 647 (PlyCB-WT-AF647). B) Probing the same samples with the GAC antibodies confirms the presence of GAC in the bands. Molecular mass markers are given in kDa.

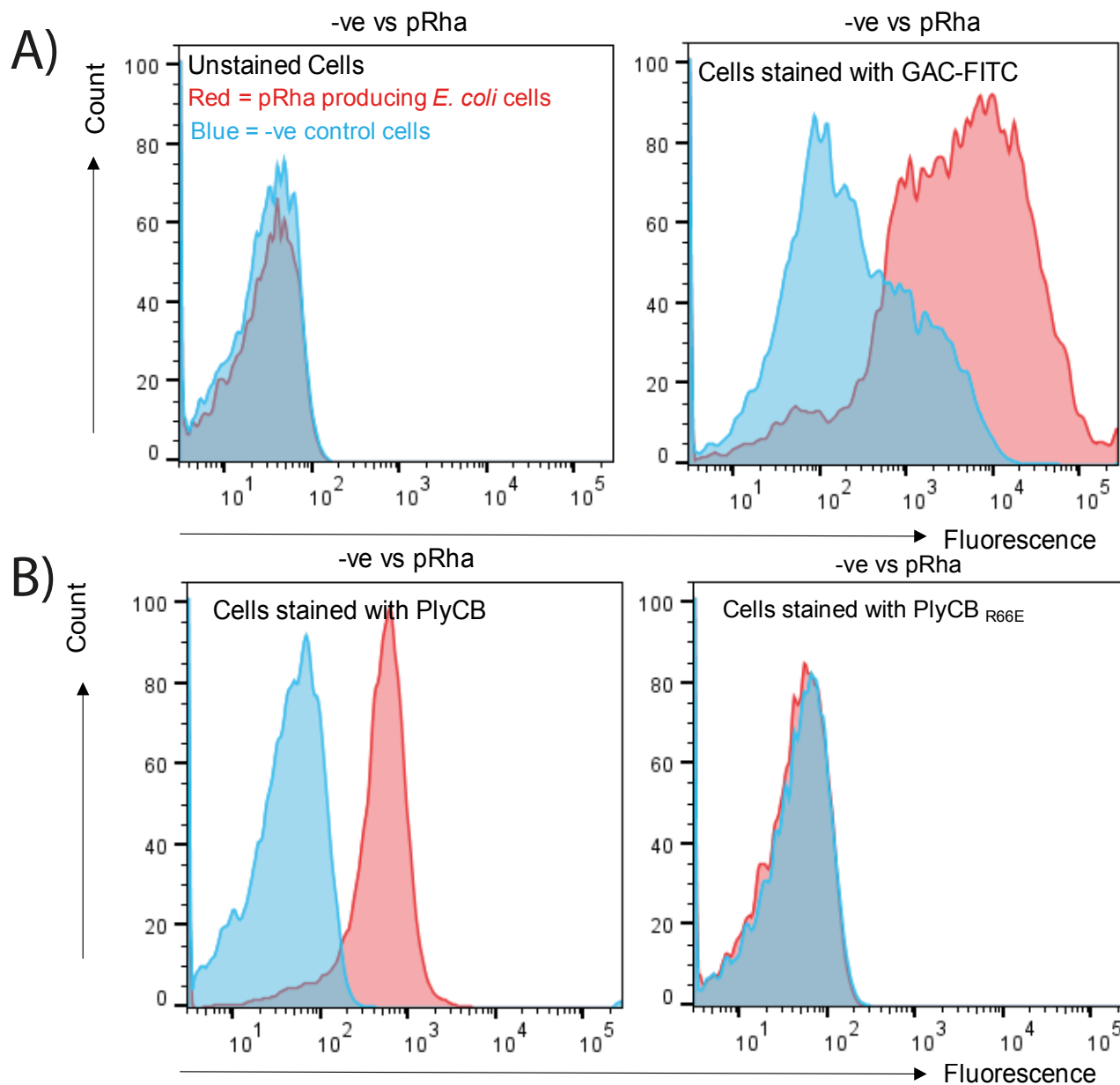


Fig 4)

PlyCB binding to *E. coli* cells were investigated by flow cytometry after labelling with PlyCB_{WT}^{AF647} and PlyCB_{R66E}^{AF645} mutant proteins. Blue: -ve control cells without pRha. Red: pRha producing *E. coli* cells. Representative histograms are shown. A) Left panel: unstained cells. Right panel: The anti-GAC antibodies (GAC-FITC) were used as a positive control to label the *E. coli* cells producing pRha. The antibodies do not bind to the *E. coli* cells carrying an empty plasmid (-ve). B) Left panel: PlyCB_{WT}^{AF647} binds to the *E. coli* cells producing pRha, but not to the *E. coli* cells carrying an empty plasmid (-ve). Right panel: PlyCB_{R66E}^{AF645} does not binds to the *E. coli* cells producing pRha.

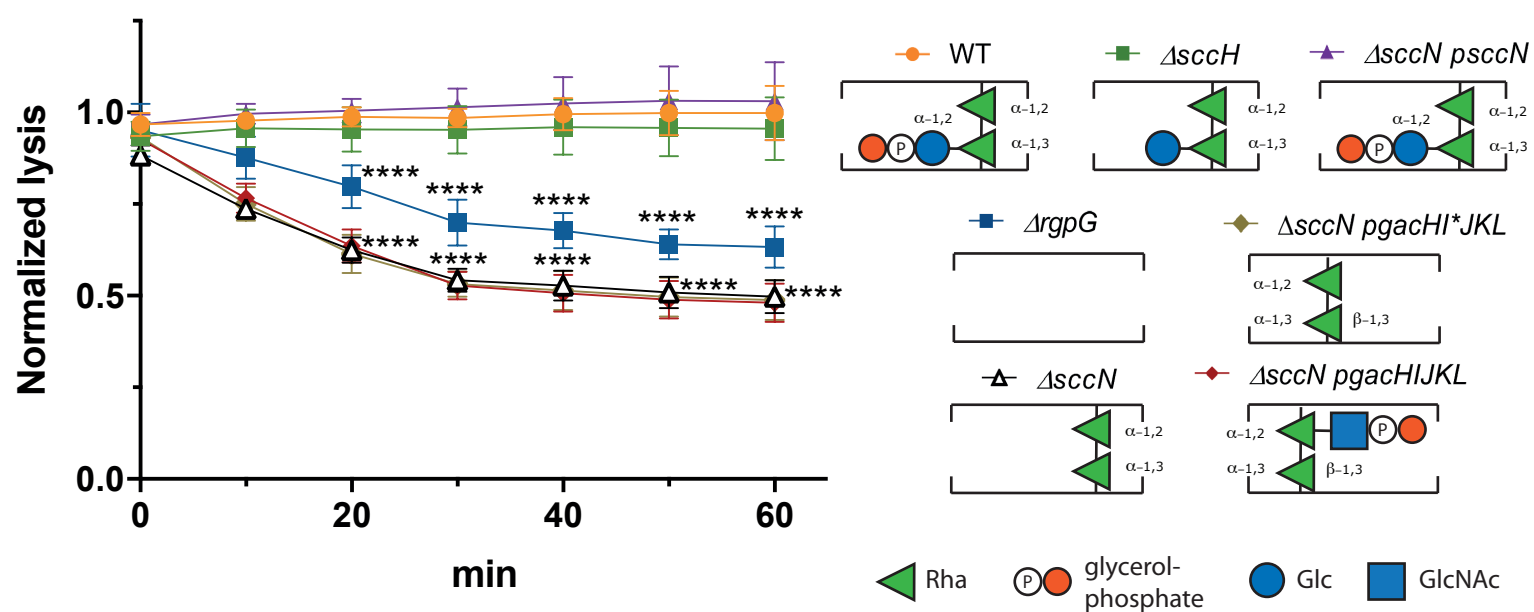
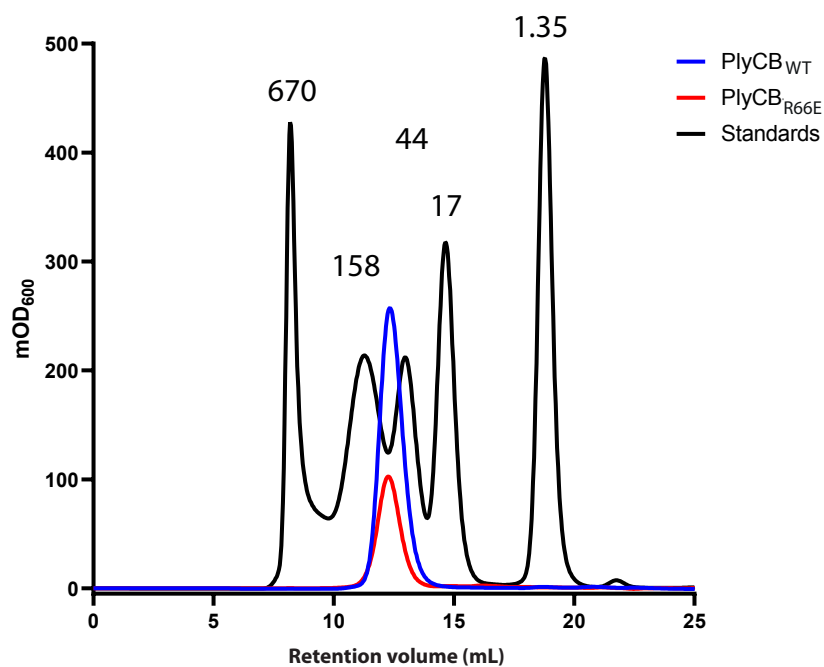


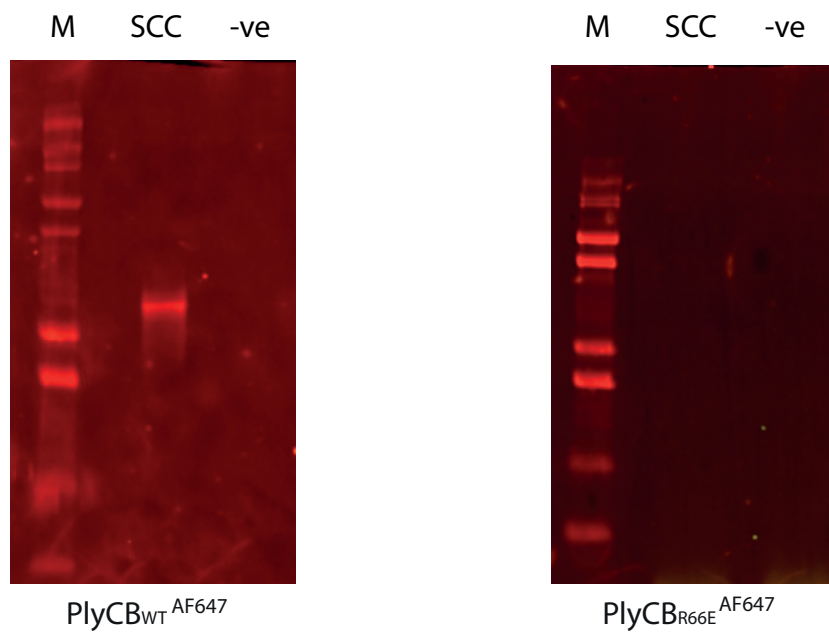
Fig 5)
The PlyC-mediated lysis of sacculi purified from *S. mutans* strains. The lysis was monitored after 10, 20, 30, 40, 50 and 60 min as a decrease in OD_{600} . Results are presented as a fold change in OD_{600} of the sacculi incubated with PlyC vs. the sacculi incubated without PlyC. Data points and error bars represent mean values of four biological replicates and standard deviation, respectively. P-values were determined by two-way ANOVA with Dunnett's multiple comparisons test.

A)



SF 1) Gel filtration purification of WT and R66E-PlyCB proteins. PlyCB_{WT} (blue) and PlyCB_{R66E} (red) proteins elute at a retention volume of 12 mL. The PlyCB octamer has a theoretical mass of 63.8 kDa, with a monomeric mass of 7.9 kDa. Gel filtration standards are shown in kDa and include bovine thyroglobulin (670 kDa, void volume), bovine gamma-globulin (158 kDa), chicken ovalbumin (44 kDa), horse myoglobin (17 kDa), and vitamin B12 (1.35 kDa).

A)



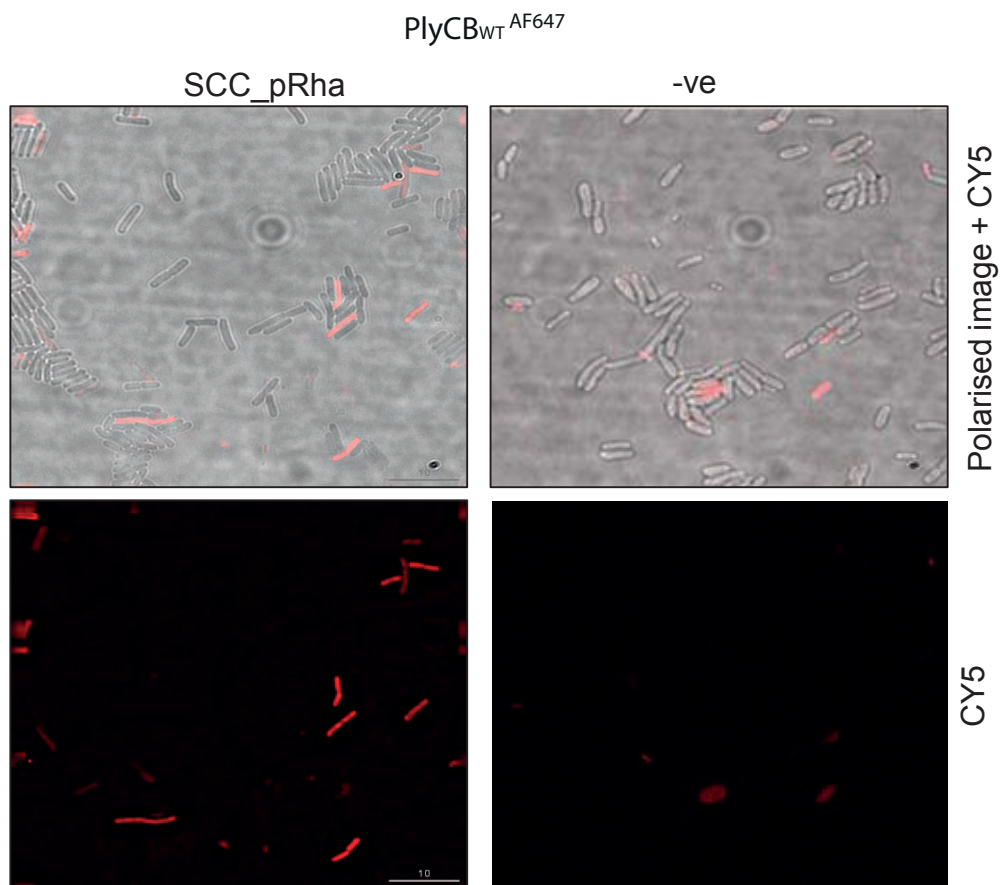
B)



SF 2)

PlyCB binds to pRha in *E. coli* cell lysate whilst PlyCB_{R66E} does not. **A)** Representative membrane blots showing that only PlyCB wildtype protein binds to total *E. coli* lysate when the pRha is expressed from the SCC gene cluster. PlyCB_{R66E} fails to bind to pRha. **B)** Control dot blots showing that both PlyCB proteins are fluorescently labelled and detected when spotted directly onto nitrocellulose membrane (1:1000 dilution - identical to Figure A).

A)



SF 3)

PlyCB binds to pRha expressing *E. coli* cells analysed using microscopic analysis: Representation of microscopic graph shows defined elongated PlyCB_{WT} stained *E. coli* bacterial cells suggesting that PlyCB binds specifically to SCC_pRha cells but not to *E. coli* cells lacking pRha expression (-ve) detected using CY5 channel.